

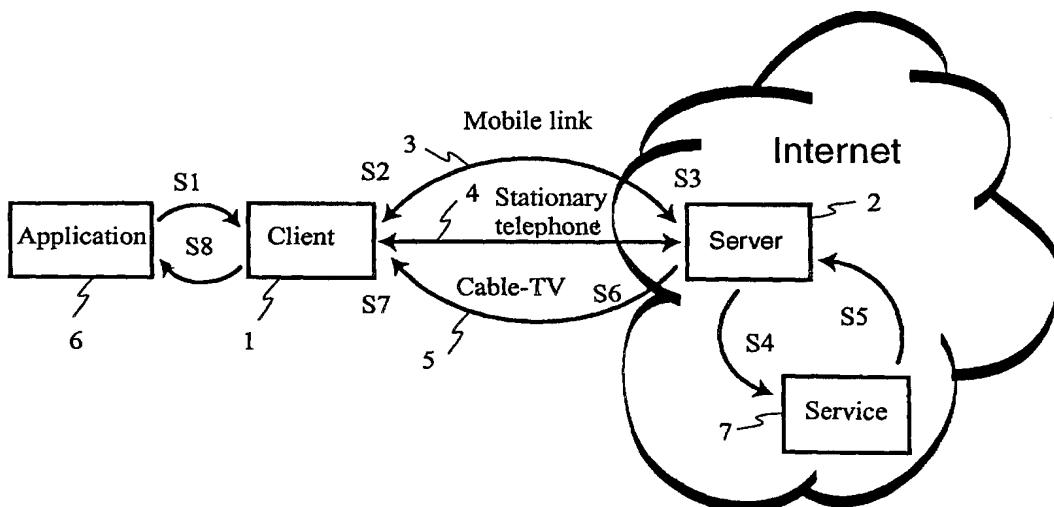


INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁷ : H04L 12/56, H04Q 11/04, H04L 29/06	A1	(11) International Publication Number: WO 00/67435
		(43) International Publication Date: 9 November 2000 (09.11.00)

(21) International Application Number: PCT/SE00/00883	(81) Designated States: AE, AG, AL, AM, AT, AT (Utility model), AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, CZ (Utility model), DE, DE (Utility model), DK, DK (Utility model), DM, DZ, EE, EE (Utility model), ES, FI, FI (Utility model), GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KR (Utility model), KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SK (Utility model), SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
(22) International Filing Date: 4 May 2000 (04.05.00)	
(30) Priority Data: 9901621-4 4 May 1999 (04.05.99) SE 0000707-0 1 March 2000 (01.03.00) SE	
(71)(72) Applicants and Inventors: KARLSSON, Mats [SE/SE]; Richtersgatan 2b-2205, S-412 81 Göteborg (SE). BERGEK, Martin [SE/SE]; Kullengatan 8b, S-412 62 Göteborg (SE). AGERVALD, Magnus [SE/SE]; N.a. Stationsgatan 99:315, S-113 64 Stockholm (SE). AXELSSON, Kristian [SE/SE]; Eklandagatan 23A, S-412 82 Göteborg (SE).	
(74) Agent: AWAPATENT AB; P.O. Box 11394, S-404 28 Stockholm (SE).	Published With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: A SYSTEM FOR DATA TRANSMISSION VIA SEVERAL COMMUNICATION ROUTES



(57) Abstract

The present invention relates to a system for transmission of data between at least two units (1, 2) adapted to inter-communicate, said system comprising several possible communication routes having different characteristics. The invention likewise concerns a corresponding method, and a terminal unit for use in the system according to the invention. More particularly, in accordance with the invention the transmission unit gathers information on the characteristics of the possible communication routes and on the basis of that information and of a given optimising condition given by the user it divides the flow of data among at least two of said communication routes whereas the receiver unit in a corresponding manner re-combines the data flow.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon			PT	Portugal		
CN	China	KR	Republic of Korea	RO	Romania		
CU	Cuba	KZ	Kazakhstan	RU	Russian Federation		
CZ	Czech Republic	LC	Saint Lucia	SD	Sudan		
DE	Germany	LI	Liechtenstein	SE	Sweden		
DK	Denmark	LK	Sri Lanka	SG	Singapore		
EE	Estonia	LR	Liberia				

A SYSTEM FOR DATA TRANSMISSION VIA SEVERAL
COMMUNICATION ROUTES

Technical Field

The present invention relates to a system for transmission of data between at least two units that are adapted to inter-communicate, said system comprising several possible communication routes having different qualities. The invention likewise comprises a corresponding method, and a terminal unit for use together with the system in accordance with the invention.

10

Background

The demands for rapid transmission of data are constantly increasing, and for that reason new solutions for data transmission are required. Existing networks intended for other first-hand uses, such as television or electricity, may be used in some cases for data transmission as well. The prospects of increased proceeds are likely to drive the development towards more extensive use of this at presently free capacity.

20

In addition, the mobile communication networks are likely to be developed to provide increasingly rapid transmission speeds. Such development will take place in several steps and will lead to multiple increases of speed on the mobile networks within the next few years.

25

In addition, more equipment will be connected wireless to peripheral equipment and other equipment, both that located in its immediate vicinity and for long-distance communication. As a result, the need for rapid and reliable data transmission increases even more.

Today's users often have access to multiple communication links to networks, such as Internet or intranets of individual enterprises. However, only one link is employed at a time, and consequently considerable capacity 5 is not used. The possibility of sending data via one communication link and of receiving data on another does exist, but in an unintelligent and static way.

In today's data transmission systems the last link to the user is the weakest one. This is true for instance 10 in the case of data services effected via GSM or stationary modem, but also ISDN and other somewhat more rapid transmission methods are bottle-necks that hold up data transmission, because the speed on Internet is many times higher.

15 Furthermore, fixed connections to Internet are installed in an increasing number of households and offices, leading to a new standard of transmission speeds available to the average user. In consequence, many future data transmission systems will have difficulties in 20 meeting the demands of the new standard, and this is true particularly as regards the mobile systems.

In addition, to achieve sufficient coverage is also a problem found in today's systems. Irrespectively of the user's whereabouts he/she will require access to Internet, and consequently facilities for Internet connection 25 must be available globally. The facilities for connection may vary from place to place and may include:

- Fixed connections
- Connection via modem
- 30 • Connection via mobile communication, and
- Connection via satellite communication.

The coverage of mobile services is restricted by the fact that the mobile networks are not fully expanded and be-

cause of problems arising from radio-shadows in the network in question, which sometimes causes coverage disappearance. No communication system is error-proof. Nor is any system able to offer capacity to an unlimited number 5 of users.

The lack of flexibility of user's choices among several different communication routes leads to poor exploitation of capacity of the respective communication systems and thus to impaired service to the individual user.

10 In addition, competition between different communication systems would be improved, were the choice between said system more flexible.

Consequently, there is a need for a data transmission system that makes use of several transmission links 15 more efficiently, both to more efficiently utilise existing transmission channels and to provide improved performance and to achieve a higher degree of user satisfaction.

From EP-O 696 010 is known a system allowing data to 20 be transmitted on several available channels. Non-available channels are made non-selectable channels by means of a notch filter system, and the application is mainly concerned with this aspect. However, this system does not solve the above problems of making data transmission operations more efficient as regards availability, coverage, costs, etcetera.

From US-5 787 079 it is known to divide a flow of data among several part channels in a radio channel, and the flow of data may be transferred in parallel on different part channels and subsequently be re-assembled by 30 the receiver. Like EP 0 699 010 this patent specification is concerned merely with the technology of transmitting data between a base station and a mobile unit. This sys-

tem does not either solve the problems with which the present invention is concerned, particularly as concerned availability, coverage and costs.

Furthermore is known a method of asymmetrical transmission, (Asymmetrical Digital Subscriber Line, ADSL), which is a method of increasing the transmission speed from Internet to the local computer whereas the speed is restricted in the opposite direction. ADSL may use a physical connection having different transmission speeds depending on the signal modulation. Also, two physical separate connections to the network may be involved. One possible configuration could be an ordinary telephone modem for the up link and the cable-television network for the down link. ADSL is, however, useable only in cases when high transmission speeds are not required in both directions. Applications such as video conferences therefore are not feasible. On account of the lack of dynamics in the system, the problems of availability, coverage and costs are not solved either.

The division of data into several parallel flows is known from other applications. Although in many cases increased transmission speeds are achieved, they fail to satisfactorily solve the problems with which the subject invention is concerned, viz. to improve the user's possibilities of selecting by himself relevant optimising conditions, such as transmission speeds, reliability, availability and costs. When parallel lines are used in existing applications normally only the same type of communication link is made use of, sometimes even the same physical link. The result is poorer performance and increased risks of total data-flow interruption.

Object of the Invention

One object of the subject invention therefore is to completely or at least partly eliminate the above-mentioned drawbacks inherent in the prior-art, and to provide a system allowing more efficient and more rapid data transmission operations.

This object is achieved in accordance with the invention as the latter is defined in the appended claims.

10 Summary of the Invention

The invention comprises a system for transmission of data between at least two units adapted to inter-communicate, said system comprising a plurality of possible communication routes exhibiting different characteristics.

15 Each unit further comprises a transmission unit which gathers information on the characteristics of the possible communication routes and, on the basis of said information and of an optimising condition given by the user, divides the flow of data, preferably in the form of a
20 data packet, among at least two of said communication routes. Each unit also comprises a corresponding reception unit, which re-combines the data flow transmitted via said at least two communication routes. An active choice of communication route thus is made to match the
25 optimising condition of the user's choice, and in addition, the choice is made on the basis of the relevant information gathered by the transmission unit. In this manner, the user is given more satisfaction, since the optimising condition is met to a higher extent than in prior-
30 art systems, while at the same time existing communication routes and the like may be made better use of by the supply of relevant status information to the transmission units. The transmission units preferably gather informa-

tion continuously and effect the division of the data flow automatically on the basis of that information, which makes the above-mentioned advantages still more apparent. Specifically suitable bits of information to be gathered is one or several of the following: fixed and variable prices for the communication routes, maximum and present speed of the latter, sizes of packets that could be transmitted, transportation time (the end-to-end transportation time for packets), ID or similar information on forwarding units, i.e. units comprised by the communication routes, and predetermined priorities made for instance by the user concerning choices of specific communication routes, specific forwarding units or the like. Based on such information it becomes possible for instance to optimise the transmission with respect to one or several optimising parameters, such as price, speed, and preferences as to intermediary forwarding units.

The system in accordance with the invention is particularly suitable for data transmission operations, wherein at least one of the communication routes is a wireless route, since the need for such active optimisation is greatest in those cases. In addition, the system is particularly well suited to solve the data-transmission problems, when one of the communicating units in turn is connected to an additional network for forwarding data to and from that network. In this manner, the transmission system in accordance with the invention may be used in some particularly sensitive parts of a transmission chain. These parts could for instance be particularly expensive, particularly slow, have a particularly small transmission capacity or the like. The invention could be used for mobile, that is wireless, connection to larger stationary networks, such as Internet.

The invention likewise concerns a method corresponding to the system outlined above, and a terminal unit to be used in such a system.

In accordance with a different aspect, the invention 5 concerns a system for transmission of data between at least two units (1, 2), which are adapted to intercommunicate, said system comprising at least two possible communication routes (3-5), each having different characteristics, and wherein each unit comprises a transmission 10 unit which is adapted to gather information on the characteristics of said possible communication routes and to divide the flow of data, preferably in the form of data packets, and which chooses one of said at least two communication routes on the basis of the information thus 15 gathered and of an optimising condition given by the user for each such data flow part. Each unit further comprises a corresponding receiver unit, which re-combines the data flow transmitted via said at least two communication routes.

20 In this manner a choice of communication route is made for each packet or the like, and thus a smooth transition is made possible from one communication route to another, while at the same time it is ensured that the preferred communication route is utilised at all times.

25 In addition, the invention concerns a corresponding method.

The invention likewise comprises a system for transmission of data between at least two units adapted to inter-communicate, said system comprising at least two possible wireless communication routes via different base 30 stations and wherein each unit comprises a transmission unit dividing the flow of data, preferably in the form of data packets, among said at least two communication

routes and wherein each unit further comprises a corresponding receiver unit, which re-combines the data flow transmitted via said at least two communication routes. Since at least some of the communication routes in this 5 case are mobile, the system becomes particularly powerful, if the networks of several different operators (base stations) are utilised. In this manner access to a larger number of frequencies is obtained, offering improved coverage, availability and transmission speed. This should 10 be compared with the use of one single network the limited frequency supply of which is to be divided among the input lines. In the above case, when the network characteristics are static and similar, gathering of information and subsequent optimisation are not, however, necessary in order to meet the objectives of the invention. 15 Because the transmission unit divides the data flow into equal-size parts and the reception unit re-combines them, the desired effect is obtained.

20 Brief Description of the Drawings

The invention will be described in the following for exemplifying purposes by means of some embodiments and with reference to the accompanying drawings, wherein:

Fig 1 is a schematic general view of the system in accordance with the invention;

Fig 2 is a schematic view of one embodiment of the system comprising several communication routes in one physical unit;

Fig 3 is a schematic representation of one embodiment of the system comprising several separate communication routes;

Fig 4 is a schematic representation of one embodiment of the system comprising a communication unit having several communication routes;

5 Fig 5 is a schematic representation of one embodiment of the system involving several users having several communication routes;

Fig 6 is a schematic view of one embodiment of the system comprising an integrated unit having several communication routes;

10 Fig 7 illustrates one example of realisation of a combined optimising operation involving several optimising parameters;

Fig 8 illustrates one example of an NDIS wrapper according to the invention including installed PAM and NDM;

15 Fig 9 illustrates one example of output signals transmitted via three NDM.s;

Fig 10 is a general view illustrating the relation between the PAM classes according to the example; and

20 Fig 11 is a general view illustrating the relation between the NDM classes according to the example;

Fig 12 illustrates schematically how different communication technologies may provide different coverage areas. The coverage will span from being nation-wide (mobile networks) to point-based (Bluetooth, LAN, modem) as 25 well as anything in between. The method described in this application makes it possible to roam between these different technologies; and

Fig 13 is a schematic representation of a system in accordance with the invention, wherein via multiple communication links the client communicates with a "remote server" via a server in the system.

Description of Preferred Embodiments

As illustrated in Fig 1, the system of the invention comprises several parallel communication links, the objective of which is to solve or reduce the problems outlined previously herein. The system comprises at least two units 1, 2, which are adapted to inter-communicate, and the system comprises a plurality of possible communication routes 3-5 exhibiting different characteristics.

Each unit comprises a transmission unit, which gathers information on the characteristics of the possible communication routes and, on the basis of that information and of an optimising condition given by the user, divides the flow of data, preferably in the form of data packets, among at least two of said communication routes. In addition, each unit comprises a corresponding reception unit, which re-combines the data flow transmitted via said at least two communication routes.

The user or the application 6 utilising the system in accordance with the invention may for instance request communication with another unit, such as to obtain a service 7 on Internet. Normally, a request of this kind is transmitted directly to the supplier of the service, who then sends his answer, if any, back via the same communication link on which the original request was transmitted. In exceptional cases the return link is not the same as the original one but in these cases the choice of up and down link is determined statically. By positioning a reception unit in the form of e.g. a computer on Internet and making all requests and answers pass therethrough active choices of communication routes within the system are made possible.

The following user case describes the flow from the instance the request is made to the return of the re-

sults. The user initially executes an operation in an application, which causes transmission of a request (step S1). The user's transmission unit, i.e. the client part of the system, processes the request and forwards it via 5 one or several communication links to the reception unit, i.e. the server part of the system, which in turn may be connected to another network, such as Internet. Together with the request is transmitted also other information needed by the server to optimise the transmission of the 10 answer to the request (step S2). The server then receives the request (step S3) and performs the requested service. This may be a local service on the server or involve the server in turn transmitting a request for a service on Internet (step S4).

15 The answer to the request is then received by the server (step S5), and based on information gathered on existing communication links the answer is divided and is returned to the client via the various communication links (step S6). As will be described in more detail below, for the optimising operation, the server uses information on the receiver, the measured transmission speed and current price information regarding the various communication routes, etcetera. The client then receives data via the different communication links and re-com- 20 bines the data flow into its original form (step S7), and the result is delivered to the application that made the 25 request (step S8).

30 The system in accordance with the invention makes an active and "intelligent" choice of different communica- tion links in both directions, resulting in possibilities of optimisation with respect to several different crite- ria.

12

In cases when the characteristics of the communication links are equal and static, it will not be necessary to optimise the data flow (steps S2 and S6 are modified).

Application 6 could be for instance a mail programme, a browser, or any other programme using data transmission. The transmission in accordance with the invention preferably is made transparent to the application because the division takes place on protocol level in the computer.

The client part 1 may be realised in the form of a driver for network connection. All communications with the network then takes place via this new driver, which consequently is able to optimise the flow. According to the invention, the client can for instance gather and thus have access to the following information, viz.:

- maximum and current transmission speeds in the respective direction, which makes speed optimisation possible by giving priority to high-speed links in case of large amounts of data.
- Type of subscription and supplying operator, rate system and current tariffs, allowing price optimisation by selection of the cheapest communication link that at the same time has sufficient transmission speed,
- identification of available forwarding servers for the different communication routes and their individual collective characteristics, which allows speed maximisation by choice of a server under low load and price minimisation by choice of a server positioned in a geographically advantageous location.

The communication routes, 3-5, consist of a connection to the receiver unit having a well defined interface to the client part. Preferably, all communication links use the same communication protocol, for example TCP/IP,

which facilitates the process of transmission of data packets via the various routes. The communication routes may consist of fixed networks, such as local data networks (LAN), electrical mains or cable-television networks, or mobile networks, such as GSM-900 (Global Service Mobile 900), GSM-1800, GPRS (General Packet Radio Service), EDGE, WCDMA, Bluetooth and Radio-LAN. When the system uses mobile networks it is likewise possible, although not obligatory, to use the mobile networks of different operators as separate communication routes. When the network of the same operator is used for several communication links the latter do not necessarily have to compete for transmission capacity, if the connection takes place for instance at different frequencies to the same base station or to different base stations. In this case the optimisation is within the specifications of the mobile networks and takes place automatically.

It is particularly advantageous to use the data transmission method in accordance with the invention, when one of the communicating units is connected to an additional network. This additional network could be an intranet, a WAN (Wide Area Network), such as Internet, or a local network (LAN) to which two or more computers are connected. The unit that is connected to the network preferably has one or several high-speed connections to the network. Requests are received and forwarded, and the corresponding answers are returned to the client. The server could for instance gather, and thus have access to the following information, viz.:

• maximum and current transmission speed in the respective directions, allowing speed optimisation by giving priority to high-speed links;

- type of subscription including operator, rate system and current tariffs, allowing price optimisation by choice of the cheapest communication link that also has sufficient transmission speed.
- 5 • identification of the client communication links and their individual and collective characteristics, allowing maximisation of speed for market segmentation and increased transmission reliability.
- 10 • user identification and password, allowing user identification and authorisation checks.

The external services available on the additional networks and with which the user is able to come into contact in this manner via the transmission system in accordance with the invention principally could be any existing and future network-supplied services. This is made possible, since all that is needed to perform the service is offered by the application as input data to the client. The input data is re-created by the client part to its original form.

20

Information Gathered by The System Units When Needed

In the system in accordance with the invention an "intelligent" division of data and re-assembly thereof takes place as well as an "intelligent" choice of different communication links to optimise the transmission with respect to different criteria. In order to achieve this it is necessary that each one of the two communicating units, i.e. the client and server parts, can send the data flow over multiple communication links and perform this intelligent division and re-assembly of the data flow. This intelligent division is made possible because the communicating units gather information allowing opti-

15

misation of the transmission of data between the units, i.e. between the client and the server.

One kind of relevant information that may be gathered is the price. By price is to be understood herein on the one hand fixed and on the other variable costs. The fixed prices could be e.g. the price per call, per connection or per month. The variable part relates to a price calculated per time unit or per packet/bit. Normally, the prices vary both in the short and the long run, a fact that the inventive system is able to take into consideration. The price information preferably is stored in the server and is spread therefrom to all clients concerned. The price quotation could be of a valid-until-further-notice character or be valid for a defined period of time only. The first type of quotation preferably is given a

a version number, offering the client program easy access to the latest quotations with only minimum communication with the server.

In addition, speed is of essence. The speed is defined herein as the amount of data transmitted per time unit. It could be speed as experienced by the user and the speed on a link included in the system. Most networks, and in particular Internet, consists of a large number of links of which the weakest one largely sets the pace of transmission between two points. In consequence hereof, the speed of the flow of received data corresponds to the speed of the weakest link. Some links will have different speeds in different directions. For example, GPRS is likely to offer larger capacity initially on the up link than on the down link.

For each type of communication link there is also defined a maximum packet size. For example, the CAN bus

uses in cars 8 bytes-size packets whereas the packet size in GPRS probably will be in the order of 300 bytes.

For each connection between two points on a network there will always be a certain minimum time for packet travel. This piece of information is interesting, since a request often could be accommodated within precisely these packets and because in those cases it is advantageous to use the link where the time of transport is the shortest. The time of transport to the client of each user is updated at a predetermined periodicity. In the case of GPRS time measuring from the server to the operator's forwarding unit (gateway, GW) is free, whereas it costs money to measure the time all the way up to the client. For that reason the transportation time to GW in this case preferably is measured so as to include a standard extra-time charge for the communication between GW and client. When the number of GWs for GPRS per country is limited, information on transportation time from a server to each operator's GW will be measurable at a predetermined periodicity and consequently it will be possible for the server to store that information and use it for all communication towards the user having GPRS units. There is nothing to prevent measurement of transportation times over the mobile links more frequently, but since this is more expensive it is desirable to avoid doing so as much as possible. Of the transmission time, the time of transport is the dominating factor in the case of small amounts of data whereas the transmission time is the decisive one in the case of large amounts of data.

Essential information could also be the identity of and the information on the servers comprised in the system. There are several reasons why users would want to employ different servers. One such reason could be for

instance a temporary stay in another country where there is a local server. By using this server, it may be possible to reduce the transmission time. Another reason may be to increase the system safety and to make it possible 5 to connect and disconnect servers without affecting the users or indeed even without the users noticing. The system also needs information on the geographical location of the servers and on the capacity of the connection to the network. Finally, information is also needed on the 10 costs for connecting the server to the network. All this information may vary with the time of day. For example, some servers may be open only during certain hours.

It may likewise be of interest to choose a link on the basis of other grounds than the price and the transmission speed. When the transmission links are three mobile links an operator may wish in the first instance to make use of the capacity of his own network before buying capacity from another operator. The priority-selection 15 information of this type may be stored in the server. In addition, during operation it could be possible for the user to make a preference whether the system is to optimise with respect to most favourable price or highest speed, as will be described in more detail in the following.

When the user's unit is being configured, some information preferably also is entered. This information relates to the type of links and the information required to communicate therewith. This comprises also information regarding how often the client updates the server and on 25 the progress of the communication.

In addition, it may be important to know the geographical location of the client. By knowing the user's position it becomes possible to select the most

advantageous server. It could also be possible for the user to enter his geographical location himself.

The table below summarises examples of information that it may be valuable for the units incorporated in the system in accordance with the invention to gather, where such information may be gathered as well as how such gathering activities may be effected.

Summarising Table on Information Gathering

Type of Info	Source of Info	Gathering Method
Price	Server	Defined
Speed (maximum)	Server	Defined
Speed (current)	Client/Server	Measured
Packet size	Server	Defined
Transportation time	Client/Server	Measured
Server ID/Info	Server	Defined
Priority	Client/Server	Defined
User Setting	Client	Defined/Measured

10

Optimisation of Data Transmission

The system is complex and there may be a considerable number of parameters and preferences that guide the optimisation procedure. However, the optimisation is based on an optimisation condition defined by the user. This optimisation condition could, however, in turn comprise several different optimising parameters. In this case different priorities should, however, be attributed to the different parameters in order to indicate their relative importance. Four examples of optimisation areas are given below.

According to the first example, the optimisation is made on the basis of the parameter price, which often is desirable. Given all necessary information on price in

19

the manner indicated above for each individual link the total price for the respective link is calculated. One of the links will be the cheapest one. Purely price-based optimisation therefore will result in all transmission 5 being carried out on the cheapest link. As the prices may vary, this may however by changed during a transmission operation.

One very important optimisation parameter is transmission speed. Speed optimisation may be carried out on 10 the basis of the following formulas:

$$f(x_1, x_2, \dots, x_n) = \max(t_1, t_2, \dots, t_n) [1]$$

$$t_i = [t_{i0} + s * x_i / v_i] * h(x_i) \text{ for } i = 1, 2, \dots, n [2]$$

$$h(x_i) = 1 \text{ for } x_i > 0 \text{ and}$$

$$= 0 \text{ for } x_i \leq 0 \text{ for } i = 1, 2, \dots, n [3]$$

$$x_1 + x_2 + \dots + x_n = 1 [4]$$

$$0 \leq x_i \leq 1 \text{ for } i = 1, 2, \dots, n [5]$$

when $n \geq 2$; where x_i is the part data transmitted on link i ; t_i the time from dispatch of the first packet to 20 reception of the last packet over link i ; t_{i0} is the transportation time between server and client in the direction in question on link i ; v_i is the speed of transmission on link i ; and s is the total amount of data to be transmitted.

25 The optimisation procedure is divided into three parts: The request is sent to the server, the server performs the service and finally the response is returned to the client. In the text below $n=3$.

In order that the dispatch of the request, in part 30 1, be as rapid as possible, it is sent on the link having the shortest transportation time, i.e. $\min(t_1, t_2, t_3)$. In-

formation on which one of the links that has the shortest transportation time is derived on the one hand from previous transmissions and on the other from the connection procedure in connection with the on-going session. Together with the request, the client also sends information on the speed (v_1, v_2, v_3) with which the client received the previous transmission on the respective link.

The time the server needs to perform the service, i.e. part 2, is beyond the control of the system and preferably is not handled by the optimisation algorithm.

The result to be dispatched from the server in part 3 to the client is to be divided among the three links of the system. During the client's request in part 1 information was sent on the client's latest reception speeds. These two pieces of information are used together with the transport time data relating to the three links in order to minimise [1] with respect to x_1, x_2, x_3 according to secondary conditions [4] and [5]. In other words, the objective is to minimise the time needed for the latest packet received, which leads to a percentage division of the flow of data between the different links.

f is a function in R^3 . The secondary condition [4] is a plane intersected in $A=(1,0,0)$, $B=(0,1,0)$ and $C=(0,0,1)$. [5] causes limitation of the defined set of f to the triangular surface that A , B , and C defines. The maximum of t_1 is in A and declines linearly towards the line BC and on line BC it discontinuously becomes zero; the maximum of t_2 is in B and declines linearly towards line AC and on line AC it discontinuously becomes zero; the maximum of t_3 is in C and sinks linearly towards line AB and on line AB it discontinuously becomes zero. In cases when the three transportation times are of corresponding magnitude an unambiguous solution x_1, x_2, x_3 is

21

obtained somewhere between A, B, and C and $t_1 = t_2 = t_3$. In case for instance $t_{10} \geq t_{30} + s / v_2$ and $t_{10} \geq t_{30} + s / v_3$, a solution is obtained somewhere on the line between B and C, i.e. $x_1=0$. In case both $t_{20} \geq t_{10} + s / v_1$ and $t_{30} \geq t_{10} + s / v_1$ a solution in point A is obtained, i.e. $x_1=1$ and $x_2=x_3=0$.

The client informs the server of the measured values of v_1 , when one of two criteria are met, viz.: either a predetermined amount of data has been received or a pre-determined time has lapsed. These two parameters are set by the user. Accordingly, the server is able to adapt the selection of communication links during rapid variations of the transmission speeds via the different links. Small amounts of data will give preference to the links showing the shortest transportation time whereas large amounts of data preferably will be sent via high-speed transmission links. Should one link cease to function, the server will be informed accordingly as a result of v_1 being lowered to zero, which de facto results in disabling of the link.

Should the client find out later that communication via this link has been resumed, the server will be informed accordingly and the link again be put into service.

As mentioned previously, it is likewise possible to combine several optimisation parameters within the scope of the invention. For instance, price and speed may be combined. In order to enable compromise solutions and thus to make it possible to offer precisely the capacity that the client is prepared to pay for, the user is offered an option to choose how the price and the speed, respectively, should be weighted in the optimisation procedure. Often, price and speed are opposites, i.e. it is not possible to simultaneously achieve both the lowest price and the highest transmission speed. The extremes of

optimisation of price and speed, respectively, have been described above. The combination thereof may be made by simple linearisation as shown in Fig 7. To the left in the drawing figure, all data are sent on the cheapest link. To the right, data is transmitted divided in the manner providing the most favourable transmission speed as calculated in the manner described above.

Optimisation may also be made on the basis of choice of server. The following procedure is followed for instance when a system consisting of three mobile links connected to Internet via several GWs is started. The client connects to the same server as at the latest instance of use. The server then sees which operator's gateway that the client communicates with. The server informs all other servers that it has a client via a certain operator's GW. All other servers then measure the transportation time to the same GW, and the result is sent to the original server. The original server sees whether another server can offer more favourable transportation time to this GW. Should a server be in a position to offer a significantly better transportation time to the operator's GW, the server hands over the responsibility to that server.

The transportation times of the servers to the most common gateways will, however, be rather static, and consequently there might be reasons for storing the corresponding data in each server.

Examples of Alternative Embodiments

The five components that are included in the embodiment of the system in accordance with the invention described above (application, client part, link, network

and server) may be combined physically in different ways, providing different areas of use and functionalities.

According to one embodiment illustrated in Fig 2 application 6 and client part 1 are integrated in one unit, 5 for example a portable computer, i.e. the drive routines for division of data etcetera are installed on the computer. Divided data are sent via cable or in wireless mode to the communication unit, which in this case is a physical unit comprising several independent telephone 10 units 1a-c. Preferably, the telephone units are located in close vicinity to one another and preferably they share the same source of power but otherwise they do not communicate with one another. The flow of data is forwarded only by the telephone units included in the communication 15 units to the network of the respective communication link. The networks of the communication links are interconnected with stationary networks, such as Internet, where the receiving server part 2 is positioned.

According to an alternative embodiment, shown in Fig 20 3, application 6 and client part 1 are interconnected in the same way as in the previous embodiment. In this case, the various communication units 1a-c are not, however, integrated in one physical unit but are separate units. If the user should want to communicate with Internet via 25 several parallel lines but is able to contribute himself with only one communication link, several links may be made available by loan of capacity from neighbouring communication units. They may have accepted the loan of capacity in advance or else the question of compensation 30 may be solved by means of an intelligent billing system. To enable the transmission between the user's client part and the various communication units to be effected as convenient as possible wireless transmission mode pref-

erably is used. The various communication links need not, however, in this case be of the same type. For example also modems, electricity mains or cable-television may be used as communication links.

5 In a further alternative shown in Fig 4 the client part 1 is combined with one or several communication units 1a-c in the same unit. In this manner an independent unit is created, which is able to communicate with a separate application 6 on one hand and towards the network of a separate communication link on the other. Thus 10 a separate communication unit is made possible, having no other function than to divide data in an intelligent manner among different lines. The application must be connected separately for instance via a serial port.

15 Application 6, client part 1 and the communication unit 1a-d could also be three physically separate units. Thus, the client part may act as the "spider in the centre of the cobweb" and on the one hand be in contact with the respective applications 6 of one or several users and 20 on the other with one or several communication links 1a-d. This situation is illustrated schematically in Fig 5. In this case the client part coordinates data from the different users and optimises the use of the different communication links. A typical field of use is by a group 25 of user, for instance consultants visiting a customer, who wish to share one and the same communication method.

Alternatively, application 6, the client part 1 and the communication unit 1a-c could also be integrated in a common unit, as shown in Fig 6. This unit thus is able to 30 satisfy all user need on the client side. Physically, the product may be similar to a small hand-held computer having several mobile telephone antennae or be a telephone that may be used as a mobile laptop modem.

The physical unit that might be used to transmit data over multiple lines thus could be realised as follows: The terminal has multiple parallel lines on one or several printed cards in a physical unit. Consequently, 5 the lines of the terminal may consist of several separate telephone units, each having its own SIM card. Alternatively, a common SIM card may be used for a number of lines. The unit is also equipped with one transmission step and one antenna per line. Also in this case it is 10 possible, as an alternative, to divide the transmission step and/or the antenna among several lines. While the lines of the terminal thus are separate they do, however, preferably share the battery and the casing. The different lines in the terminal preferably are also able to inter-communicate internally as regards certain information, such as control information, such that all lines do 15 not transmit on maximum efficiency simultaneously, which might be beyond the battery capacity. Also, the different lines may share the same bluetooth, buffer memory, etcetera. The various lines of the terminal may consist of 20 the same kind of units or else they could be different ones, such as three GPRS units, two GPRS units and one EDGE unit, one GPRS unit, one EDGE unit and 1 W-CDMA unit.

25 The terminal may be conceived to transmit data only, but could also have conventional mobile-telephone functions, i.e. it may be used as an ordinary mobile telephone. The terminal having multiple lines could also comprise the function of a hand-held computer, i.e. the computer and the terminal built into one unit.

In some applications there is a need for higher-speed transmissions than conventional network interface cards could cope with. To obtain higher speeds it becomes

necessary to buy more advanced network interface cards at many times higher costs. The invention makes it possible to solve this problem at a fraction of the cost by transmitting data on several parallel lines when the cards are 5 standard network interface cards.

Examples of Implementation of the Invention

In the following an embodiment will be described, using several transmission lines in parallel. More specifically the embodiment is directed to a packet-based parallel data transmission based on a client/server solution, where the operative system, here the network subsystem of Windows NT, is modified in such a way that the TCP/IP-stack is able to transparently send packets over 10 multiple transmission lines. The same process is repeated for incoming traffic, i.e. the TCP/IP-stack is not aware 15 of the fact that the packets are transmitted in parallel.

In this embodiment Windows NT is used as the operative system of the system. The subsystem of Windows responsible for dealing with the lower layers of networking 20 is called the *NDIS wrapper* (NDIS = Network Device Interface Specification) and provides the interface between the various networking parts. The parts themselves are represented as *Miniports*, usually in the form of files 25 ending with the suffix "SYS". These device drivers can be divided into two categories, *protocols* and *adapters*. The NDIS wrapper is responsible for loading these drivers and for keeping track of which protocols are using which adapters. This is also known as *binding*, i.e. a protocol 30 binds to one or several adapters.

The protocol device drivers implement a specific communication protocol at the transport layer, e.g. IP. They are responsible for sequencing a data stream into

packets and framing these packets. By framing is here meant the process of adding header information to the packet.

The adapter device drivers are responsible for
5 transmitting blocks of data to the physical medium. They
are also responsible for signalling to the upper layers
when incoming data is available.

The TCP/IP stack on Windows NT is the component responsible for implementing the IP protocol and associated
10 services, e.g. TCP and UDP. The protocol itself requires an IP address for each adapter it binds to. The stack is also capable of *IP forwarding*, meaning that any IP packets arriving with a destination address other than any of the local adapters' will be "forwarded" to another computer.
15 The TCP/IP stack decides which adapter to use when sending IP packets based on the contents of a *routing table*. This table contains a list of IP addresses and associated *masks*, i.e. values that specify which parts of the corresponding IP address should be considered valid when
20 doing routing calculations.

It is preferred that the system is as transparent as possible. This could be accomplished by creating a so-called *Intermediate NDIS Driver*. Such an intermediate driver is a driver arranged between existing protocols
25 and adapters. They appear as adapters to protocols, and as protocols to adapters. The intermediate driver according to the invention is called a *Protocol Adapter Module*, (PAM). The PAM utilizes adapters known as *Network Driver Modules* (NDM) when communicating with the transmission
30 links. A NDIS wrapper with installed PAM and NDM is illustrated in Fig 8.

The PAM binds to the appropriate adapters and/or NDMs. It then exposes itself as a network adapter and is

subsequently bound to the overlying protocol(s). It modifies and, if necessary, splits the outgoing packets and adds certain header data to the split packets (also called *fragments*). It keeps track of which NDMs are active and what their transmission performances are. It then uses this as a basis when sending the fragments. It also queues and, if necessary, reassembles incoming fragments before passing them on to the bound overlying protocols.

There is preferably one NDM for each logical link. These are either standard Miniports provided by the manufacturer of the hardware, or specially developed if no regular Miniport is available. In either case they are preferably implemented as standard NDIS adapters.

All device drivers could be implemented as dynamic link libraries (DLLs).

In a preferred embodiment the NDIS will call the drivers during the binding phase when it analyses the registry. The adapters will receive a single call to their initialization routine, at which time they should attempt to claim hardware resources and contact their NICs (Network Interface Card). Protocols, on the other hand, will receive a call for each adapter they're bound to. This allows them to allocate a memory area in which they can store state information on a per-adapter basis.

Adapters will be queried by NDIS as to their capabilities, e.g. maximum packet size, media supported, link speed, etc. NDIS will use this information when processing packets.

When a protocol sends a packet through NDIS, it is expected that the adapter responsible for the actual transmission signals the protocol when it has finished sending the packet. Calling a function in NDIS with a

29

pointer to the packet in question carries this out. This process is known as *completing* the packet.

When the TCP/IP stack has finished opening and initializing its adapters, it broadcasts an ARP (Address Resolution Protocol) Request with the IP address being that of the local interface(s). It does this in order to find out if there are any conflicts in the reachable networks. If no reply is received, the stack assumes that there is no conflict and activates the adapter. If, on the other hand, an ARP reply is received, there is obviously a conflict, so TCP/IP prints an error message and disables the adapter.

When TCP/IP sends a packet it needs to know two things: which adapter to use and which Ethernet address to use for the destination address. The Ethernet address is retrieved by means of an ARP request, and is stored in a table. The adapter to use is resolved by consulting the routing table.

An entry in the routing table preferably contains the following information:

- Network destination: This is the destination IP address.
- Netmask: Before making any comparisons based on the IP address, a logical AND between it and the netmask value is first performed. This allows for making comparisons based on a group of values instead of one at the time.
- Gateway: This is the IP address of the "gateway", i.e. the host that knows where to send packets with an IP address matching that of the Network destination.
- Interface: The Interface is the IP address of the adapter responsible for sending the packet.

- Metric: The Metric is a simple value that would be used when choosing which Interface to use if more than one was possible for the actual route.

5 *The Transmission Process*

The system according to the example basically consists of two bi-directional data flows; one outgoing from the overlying protocol, and one incoming from the underlying NDMs. However, several incoming or outgoing flows 10 could be used as well. The data flows are received and sent as packets. Each incoming packet is split into fragments if it is too large to fit within a single NDM packet. The fragment (which can consist of a single packet if the packet is small enough) has header data 15 added to it before it is sent to an NDM for further transmission. This is illustrated schematically in Fig 9.

The process is repeated in reverse for incoming fragments, except that multi-fragment packets will be held in a list until all fragments belonging to a certain 20 packet has arrived. If a certain amount of time passes without any new fragments belonging to the partially re-assembled packet have arrived, the packet is considered lost and all fragments belonging to it are discarded.

25 *An Example of an Inventive Protocol Adapter Module (PAM) according to the invention*

The PAM is responsible for the "one-to-many" interface, i.e. it poses as a single Ethernet adapter to NDIS, but binds to several adapters below it. A PAM is used by 30 the client computer, but also by the router gateway. The only difference between them is that the PAM on the router gateway binds to an Ethernet adapter as well as to the Serial NDM, and intercepts and replies to ARP re-

quests regarding the client's IP address coming from other computers than the router gateway.

The PAM according to the invention preferably comprises the following classes of objects. The functional 5 relationship between these classes in the PAM is illustrated in Fig 10.

- *Main*: This is the class responsible for the creation of all other objects. It also contains the global driver entry function called by NDIS as soon as the 10 device driver is loaded. The Main object creates all other objects and stores their addresses in public pointers. The other objects are able to access each other through theses pointers by using a global pointer to the Main object.
- *SplitJoin*: This class is the "spider in the web". It receives outgoing packets from UpperAdapter and, if necessary, splits these packets into smaller parts called *fragments*. It then lets Framing add fragment-specific headers before calling the LowerAdapter decided by Router and asking it to send the fragment. It 15 also lets Arp construct fake ARP replies and returns them to UpperAdapter.

The process is repeated in reverse for incoming fragments; Framing verifies them and removes the unnecessary header data. The fragments are then stored in a list, waiting for the other fragments that make up the original packet to arrive. When the whole packet has been received, SplitJoin reconstructs it and lets 20 Framing add an Ethernet header before sending it to UpperAdapter.

- *Router*: This class decides which NDM will get the next packet fragment. It uses data from the Statistics ob-

jects and compares transmission speeds, cost, etc. As of now the algorithm simply assigns the first packet to the first NDM, the second packet to the second NDM, etc in a round-robin fashion.

- 5 • *Framing*: The Framing class provides functionality used when adding fragment headers to the packets handed to it by SplitJoin. It also verifies the integrity of incoming fragments and removes the headers. When SplitJoin sends incoming packets to the TCP/IP stack, Framing adds the fake Ethernet headers. In order to do this, it queries Config as to what the PAM's fake Ethernet address is.
- 10 • *Arp*: This class provides functionality used when detecting and replying Ethernet ARP requests. Since TCP/IP will not send anything until it has received an Ethernet address corresponding to the IP destination address, a "fake" Ethernet address must be provided. At the same time, it is not desirable to provide fake Ethernet addresses for anything but the client's IP.
- 15 Thus Arp must query Config what the client's IP address is.
- 20 • *UpperAdapter*: This is the class providing the "one" interface. It is responsible for the upper communication with the protocols bound to the PAM, e.g. TCP/IP.
- 25 It informs NDIS that it is a "late binding" adapter, i.e. it will not register with NDIS as an Ethernet adapter until at least one NDM has been bound to the protocol part of the PAM.
- 30 UpperAdapter is also responsible for handling any queries from NDIS, such as the PAM's Ethernet address, link speed, maximum packet size supported, etc.

- *LowerAdapter*: This class represents the lower adapters bound to the PAM. Since it is the protocol's responsibility to keep state information for its adapters, Protocol creates one instance of this class for each NDM that binds. Their main responsibility is notifying SplitJoin of incoming packets.

In the modified PAM on the router gateway computer in the Prototype Test Configuration, a special LowerAdapter is created for the Ethernet card used when doing IP-forwarding. This LowerAdapter's only task is to intercept ARP requests from other computers regarding the client's IP. Since ARP requests aren't forwarded to the client, the PAM creates an ARP reply and returns it without it even reaching the TCP/IP layer.

- *Protocol*: This class represents the protocol portion of the PAM. During the binding phase, it will be notified by NDIS when it should open the adapters bound to it. As soon as it has opened the first NDM, it calls UpperAdapter and informs it that it can expose its interface to NDIS.
- *Statistics*: There is one Statistics object for each LowerAdapter, and they have a one-to-one relationship. The Statistics object is called by its LowerAdapter partner each time it sends or receives a packet. It is also informed of bad or incomplete packets, etc. The Router object queries the Statistics objects when making its decision of which LowerAdapter will be used when sending the next packet.
- *Control*: This is the class responsible for communication with other processes. It does so by responding to *Device IO Control Codes*, or *IOCTLs* (Device Input Output Control Code). These are sent by e.g. the user in-

terface when it requests that a transmission parameter be changed.

- *Config:* UpperAdapter calls this class when it initializes. This allows it to read any PAM-specific parameters stored in the registry. It is also called by the LowerAdapters and reads parameters stored on a per-NDM basis. Such parameters could be COM-port to use, 5 transmission speed, inactive at start, etc.

10 An example of an Inventive Network Driver Module (NDM) according to the invention

The NDM preferably appears as an Ethernet adapter to NDIS. The NDM according to the invention comprises the following classes of objects. These classes and the relationship between them are illustrated in Fig 11. 15

- *Main:* This is the class responsible for the creation of all other objects. It also contains the driver entry function. The only global symbol in the NDM is a pointer to the Main object. It, in turn, has public pointers to all instances of the other objects. Main does not participate in any actions once the other objects have been initialized; it merely acts as a "global" data area. 20
- *Adapter:* This class provides the NDIS adapter interface callbacks. It opens the underlying serial port through use of Serial's Open method. Adapter also contains callbacks called by Serial when it notifies of new incoming packets or the completed sending of an outgoing packet. Sent packets are then completed to 25 NDIS.
- *Arp:* This class analyses packets received from the overlying protocol. If it turns out to be an Ethernet ARP Request, Arp provides methods for constructing a 30

corresponding ARP Reply. Since the PAM intercepts any ARP Requests before they reach the NDM, the functionality of this class is not used when the NDM is bound to the PAM only.

- 5 • *Serial*: This class is directly responsible for the translation and transmission of the packet data over the RS-232 medium. It first sends the Serial Test Protocol header, followed by the packet data itself. This send-procedure is carried out in an asynchronous manner. It also receives (and, if synchronization is lost, resynchronizes) data transmitted in the opposite direction. This data is then passed on as a chunk to the callback specified when initializing the Serial class through use of the Open method.

15 Data sent to or received from the IO subsystem is transferred by means of *I/O Request Packets* (IRPs), a data structure used by all kernel components in Windows NT.

20 The implementations for the PAM and the NDM presented above are however merely an example, and several alternatives are possible. For example, the implementation above is based on Windows NT as the operative system (OS). However, the invention may as well be implemented on other OS, such as UNIX/Linux. The references given above to Windows should therefore not be construed as only meaning the specific features and objects provided by Windows, but to comprise any similar feature and object in other OS.

Further, one additional interesting application of the system according to the invention in the mobile data area is the combination of circuit-switched and packet-

switched data links. One can use a regular GSM (Global System for Mobile communication) link, or more likely a High-Speed Circuit-Switched Data. Evolution of GSM that allows higher transfer speed HSCSD (High-Speed Circuit-Switched Data) link and combine that with a GPRS (General Packet Radio Service) link. The HSCSD link would terminate as usual, though fragments sent through it would be directed to the router gateway, since the HSCSD connection eventually ends up on the Internet. The GPRS link, being packet-switched from the beginning, would be treated in a similar manner. The greatest point of having this setup, would be that the GPRS link could provide an "always-on" quality, enabling the user to be connected constantly without paying except when transferring data.

The HSCSD link could then be enabled when the need for more bandwidth arises, keeping the total cost at a minimum while at the same time providing greater transmission speeds than is possible with either GPRS or HSCSD alone. There are a number of possible optimization options, such as the balancing of speed and cost, keeping the delay introduced when connecting with the HSCSD link at a minimum, etc.

An Example of Another Inventive Embodiment of the invention

The invention, as described above, may further be used as a method by which a computer may transparently use multiple communication links with individual IP addresses. The method enables roaming between IP network technologies.

The benefits of this solution are many:

- The connection is maintained even when individual connection links are dropped because of coverage, over-load, bad reception or any other reason
- Only one IP address is used externally which removes any troubles when changing between the links (only for IP addressing)
- It is possible to optimise the transfer for speed or cost
- Roaming between different technologies or different instances of the same technology is possible (e.g. it is possible to roam with WLAN which in itself has problems providing the possibility). During the time the user is not covered by WLAN, any other communication link is used to maintain the connection.
- The user will get the benefits of all the different communication links, giving a solution that is better than any of the individual links.

In the connected world of today, having an Internet connection is for many people and applications an absolute necessity. Until now people have connected to the Internet through a multitude of methods, e.g.

- Modem
- Local Area Network
- ISDN
- Cable
- Etc

So far only one of these has been the active communication link at any given time.

Although ISDN is in fact able to use multiple communication links (B-channels) in parallel, those links terminate at the same location and are circuit switched. In

principle, therefore, ISDN may be regarded as a single communication link.

Mobility is an issue that has not been of great importance in the past. Some advances have been made but
5 the available bandwidth has been much too low to render it of any practical use. Mobile communication solutions have mostly been of a semi-stationary character, e.g. portable computers with a built-in modem that is used to connect to the Public Switched Telephone Network (PSTN).

10 In the future we are likely to see a more complex scenario with a multitude of optional ways to connect to the Internet. Depending on various reasons (required bandwidth, cost issues, coverage, practical feasibility etc) the user will normally select one of the available
15 communication links at a time. This will work for stationary users who can make a selection and continue using it for a few hours. For mobile users, the situation is not quite as simple, even though increases in mobile bandwidth will make mobile data communication more interesting.
20

Moving around, the user will be within coverage of different network access methods at different times, as is shown schematically in Fig 12. At home or in the office, wireless LAN (WLAN) will probably be the preferred
25 method. At other times the user may be able to use connection points based on Bluetooth or other short-range systems. When out of range of all other systems a mobile communication system such as GPRS, EDGE or WCDMA (Wide-band Code Division Multiple Access) or any of their successors will provide coverage. Of course more traditional methods like LAN or modem are still possible. The following table indicates the range of some known communication technologies:

Long-range (A)	Medium-range (B)	Short-range (C)
GSM	WLAN	56k modem
HSCSD		ISDN
GPRS		LAN
EDGE		Bluetooth-LAN
WCDMA		
PCS		

- Switching between these systems means that the user will constantly change the IP address by which it is known on the network. This, in turns, means that the user will have to re-establish connections. This involves:
- Disconnecting from the first communication link
 - Connecting via a new connection link (which may or may not function)
 - Restarting transfers that are already in progress (file transfers etc), thereby aborting transfers that may have taken minutes or even hours so far but have not finished
 - Re-establishing IPSec tunnels since the authentication will fail because the IP address has changed

To some extent WCDMA and other wide spread mobile systems will make it possible to be mobile while transferring data. The cost associated with such methods will make it infeasible for constant use. The user will therefore still have to choose the connection method.

A solution to the problem described above is to tunnel the data on multiple communication links. One or several of these links are active at any one time and the system will support adding or removing links, either

40

manually or automatically. The reason that communication links will be added or removed could be, but is not restricted to:

- Losing coverage
 - 5 • A communication link that is faster or cheaper than the one(s) currently used has been added (e.g. arriving to the office)
 - Manual selection by the user
- 10 To enable the solution, two functional entities must be added:
 - Client software
 - Server software
- 15 The core of the client software is placed below the TCP/IP stack of the client. The client software is responsible for forwarding data on one or more of the available communication links. An embodiment of the system according to this aspect of the invention is schematically illustrated in Fig 13.
- 20

However, the client is not restricted to use three communication links as depicted in the illustration. Also, the type of link is independent of the solution as long as all the communication links are capable of transporting IP packets to and from the Internet.

25 The server software receives data from multiple clients, each of which may transmit data over multiple communication links. The server forwards the data to the remote server and relays the response back to the client that made the initial request. The remote server may be any server supplying information to the client, e.g. a web server.

The system may offer the possibility to restrict the usage to individual users or groups of users.

As data is received from remote servers by the server there are two ways to identify which client should 5 receive the data:

- Port numbers
- IP addresses

These two methods will be described in the following 10 two sections.

The benefits of this solution are many:

- The connection is maintained even when individual connection links are dropped because of coverage, over-load, bad reception or any other reason
- Only one IP address is used externally which removes any troubles when changing between the links (only for IP addressing)
- It is possible to optimise the transfer for speed or cost
- Roaming between different technologies or different instances of the same technology is possible (e.g. it is possible to roam with WLAN which in itself has problems providing the possibility). During the time the user is not covered by WLAN (Wireless Local Area Network), any other communication link is used to maintain the connection.
- The user will get the benefits of all the different communication links, giving a solution that is better than any of the individual links.

30

Port addressing works by using port numbers to identify the intended destination of data. Data sent from a

client will have one of the IP addresses of the client as sender. The server will open a connection to the remote server, replace the sender address with its own (IP_{S0}) and then forward the data. When the server receives a reply 5 it examines the port address to decide which client should receive the result. The server then sends the data, together with any other information the client needs to reconstruct the IP packet.

The PAM (Protocol Adapter Module) reports an IP address to the TCP/IP stack. This address may be the IP address of one of the links but it does not have to be. If 10 it is the IP address of one of the links, the address may still be used even if that communication link is dropped, thus making it possible to maintain a connection even if 15 the link with that IP address is dropped.

Port addressing works for most transfers. It specifically works for web (HTTP) transfers. Port addressing will be important when IP addresses are a scarce resource since a single IP address (IP_{S0}) on the server can serve 20 many clients. More clients can be served through this method by allocating additional IP addresses from the pool of addresses (IP_{S0} - IP_{Sn}) for port addressing.

In some cases the port addressing may not be used.
In those cases IP addressing will be used.

25 The server is in this case given a range of IP addresses (IP_{S0} - IP_{Sn} in the illustration, minus any that are used for port addressing) that may be given to clients. These addresses are allocated to clients on a temporary basis, in much the same way as when using DHCP (Dynamic 30 Host Configuration Protocol).

The client (or more specifically the PAM) requests a dynamic or static IP address from the server. This address (IP_{Sx}) could be the one that PAM reports to the

TCP/IP stack, although the stack could be configured to use a static IP address that will never be visible outside the client. All packets sent will have this address as the sender address. PAM sends packets (after possibly 5 having split them into smaller fragments) on the multiple communication links, each with its own IP address, through gateways to the server. The server reassembles the packets and forwards the resulting IP packet, i.e. a packet identical to the one that was sent from the IP 10 layer of the Client. The IP packet that leaves the server has the IP address IP_{sx} as sender.

IP packets sent from the network to the client are addressed to the IP address, belonging to the client (IP_{sx}). Packets sent to that address are routed to the 15 server, which forwards them to the client by changing the destination IP address on the fly, possible after having split the packet into fragments. The server does not have to use port numbers since the IP address uniquely identifies the destination.

20 PAM recollects the fragments and reconstructs the initial IP packet, which is passed to the IP layer of the TCP/IP stack.

The client sees PAM as a regular network adapter. The network likewise views the IP address of the client 25 as the only IP address of the client. What happens in the tunnel is completely transparent to the communication of the above layers. Consequently, all applications will work as intended, ignorant of the multiple communication links used to transport the data.

30 IP addressing requires the use of an extra IP address for every active client. Because of that fact port addressing will be used whenever possible as long as IP addresses are scarce.

The invention may be used to combine different known communication methods, such as HSCSD & GPRS. In this case the user has an HSCSD (High Speed Circuit Switched Data) card in the computer and an external GPRS telephone. Each 5 of these could be used for mobile data communication. With the solution presented in this paper, they can be used in conjunction. GPRS will then be used to stay online while HSCSD provides the possibility to give dedicated bandwidth. To the user this means less cost and 10 higher bandwidth. To the operator it means a more efficient use of radio resources.

Another example is to combine WLAN & WCDMA. In this case the user has a computer with WLAN and WCDMA capabilities, either internally or externally. WLAN is the 15 preferred method of communicating since it provides higher bandwidth as well as lower cost. WLAN, however, lacks coverage and may not provide the possibility to roam. WCDMA is therefore used to maintain the connection when moving between different WLAN areas. It is not necessary to transfer large amounts of data over WCDMA during those times. Only small amounts of data need to be transferred to keep the connection alive until the user reaches another area that is covered by WLAN. As soon as 20 the user once again is within WLAN coverage, either WLAN or WCDMA is used to report the new IP address for the 25 WLAN link to the server.

Still further, Bluetooth and GPRS/EDGE/WCDMA may be combined. In this case a mobile communicator is equipped with both Bluetooth as well as the capability to communicate with a mobile communication network such as GPRS, 30 EDGE or WCDMA. While being within range of a network that is accessible via Bluetooth that network is the primary choice because of higher speed and lower cost. When mov-

ing around, the mobile network is used to transfer data. By implementing the method described in this paper the communication equipment is always on-line and always using the preferred technology among those which are available.

The invention has been described herein by means of one embodiment. Several modifications are, however, possible. For instance, the system may be used for transmission between two equivalent units, none of which operates as server and is connected to an additional network. In addition, other types of optimisation are conceivable, and consequently also other types of information are required. In addition, the information could be gathered continuously at predetermined times/or when needed, or else be entered at separate occasions. The system could also comprise a larger number of units, such as more clients who communicate with one common server. These and other obvious varieties must be regarded to be within the scope of protection of the invention as the latter is defined in the appended claims.

CLAIMS

1. A system for transmission of data between at least two units (1, 2) adapted to inter-communicate, said system comprising at least two possible communication routes (3 -5) having different characteristics, and wherein each unit comprises a transmission unit which is arranged to gather information on the characteristics of the possible communication routes and, on the basis of said information and of an optimising condition given by the user, divides the flow of data, preferably in the form of data packets, among at least two of said communication routes, and wherein each unit further comprises a corresponding receiver unit, which re-combines the data flow transmitted via said at least two communication routes.

2. A system for transmission of data between at least two units (1, 2) adapted to inter-communicate, said system comprising at least two possible communication routes (3 -5) having different characteristics, and wherein each unit comprises a transmission unit which is arranged to gather information on the characteristics of the possible communication routes and to divide the flow of data, preferably in the form of data packets, and which, for each such divided part on the basis of said gathered information and of an optimising condition given by the user, chooses one of said at least two communication routes, and wherein each unit further comprises a corresponding receiver unit, which re-combines the data flow transmitted via said at least two communication routes.

3. A system as claimed in claim 1 or 2, wherein said transmission unit is adapted to continuously gather information on the characteristics of said communication routes (3 -5) and to automatically perform said division 5 of the flow of data on the basis of that information.

4. A system as claimed in any one of claims 1-3, wherein the characteristics on which said transmission unit gathers information are one or several of the following ones, viz. fixed and variable prices, maximum and 10 current speed, packet size, transportation time, ID or similar information on forwarding units, and predetermined priorities.

5. A system as claimed in any one of the preceding claims, wherein the optimisation condition comprises one 15 or several optimising parameters, such as price, speed, preferences with respect to intermediary forwarding units.

6. A system as claimed in claim 5, wherein the optimising condition comprises several optimising parameters 20 and an indication, given in advance, of priorities relating to said optimising parameters.

7. A system as claimed in any one of the preceding claims, wherein at least one of said communication routes (3) is a wireless route.

25 8. A system as claimed in claim 7, wherein said communication routes are wireless communication routes comprising several base stations.

9. A system as claimed in any one of the preceding claims, wherein at least one of said communicating units 30 (2) in turn is connected to a further network for forwarding data to and from said network.

10. A system as claimed in any one of the preceding claims, said transmission unit comprising a module having

a uniform interface to superior systems while at the same time being capable of communicating via said at least two communication routes, the latter preferably differing with respect to at least interface or transportation medium.

11. A method of transmitting data between two units (1, 2) adapted to intercommunicate via at least two possible communication routes (3 -5) having different characteristics, comprising the steps of:

10 in said transmission unit gathering information on the characteristics of the possible communication routes;

in said transmission unit further gathering information on an optimising condition given by the user;

15 in said transmission unit dividing said flow of data on the basis of said information;

transmitting the thus divided data, preferably in the form of data packets, to said at least two communication routes; and

20 in said receiver unit receiving and re-combining the data flow transmitted via said at least two communication routes.

12. A method of transmitting data between at least two units (1, 2), which are adapted to inter-communicate via at least two possible communication routes (3, 5) 25 having different characteristics, comprising the steps of:

in said transmission unit gathering information on the characteristics of the possible communication routes;

30 in said transmission unit further gathering information on an optimising condition given by the user;

in said transmission unit dividing said flow of data on the basis of said information;

transmitting each part of the thus divided data flow, preferably in the form of data packets, to one of said two communication routes, the choice of communication route for each part being controlled by the information gathered and by the optimisation condition given; and

in said receiver unit receiving and re-combining the data flow transmitted via said at least two communication routes.

10 13. A method as claimed in claim 11 or 12, wherein the characteristics on which said transmission unit gathers information are one or several of the following: fixed and variable prices, maximum and current speed, packet size, transportation time, ID or similar information on forwarding units, and predetermined priorities.

14. A method as claimed in any one of claims 11-13, wherein the optimising condition comprises one or several optimising parameters, such as price, speed, preferences regarding intermediary, furthering units.

20 15. A method as claimed in any one of claims 11-14, comprising the further step of transmitting the data received from the receiver unit and forwarding it to another network.

16. A terminal unit for transmission of data via at least two communication routes (3 -5), said terminal unit comprising a transmission unit (1) which is connected to said at least two communication routes (3 -5) and which is adapted to divide a flow of data into parts and to transmit said data flow parts essentially simultaneously on said at least two communication routes (3 -5), and a receiver unit (2), which is adapted to receive parts of a data flow on different lines and to re-combine them into one flow.

17. A terminal unit as claimed in claim 16, wherein at least one and preferably all communication routes (3) are wireless routes.

18. A terminal unit as claimed in claim 17, wherein 5 it comprises several antennae and preferably one antenna per wireless communication route.

19. A terminal unit as claimed in claim 17 or 18, said terminal unit comprising several transmission steps, and preferably one transmission step per wireless communication route. 10

20. A system for transmission of data between at least two units (1, 2) adapted to inter-communicate, said system comprising at least two possible wireless communication routes (3) extending between different base stations, and wherein each unit comprises a transmission unit which is arranged to divide the flow of data, preferably in the form of a data packet, among at least said two communication routes (3), and wherein each unit further comprises a corresponding receiver unit, which 15 combines the data flow transmitted via said at least two communication routes. 20

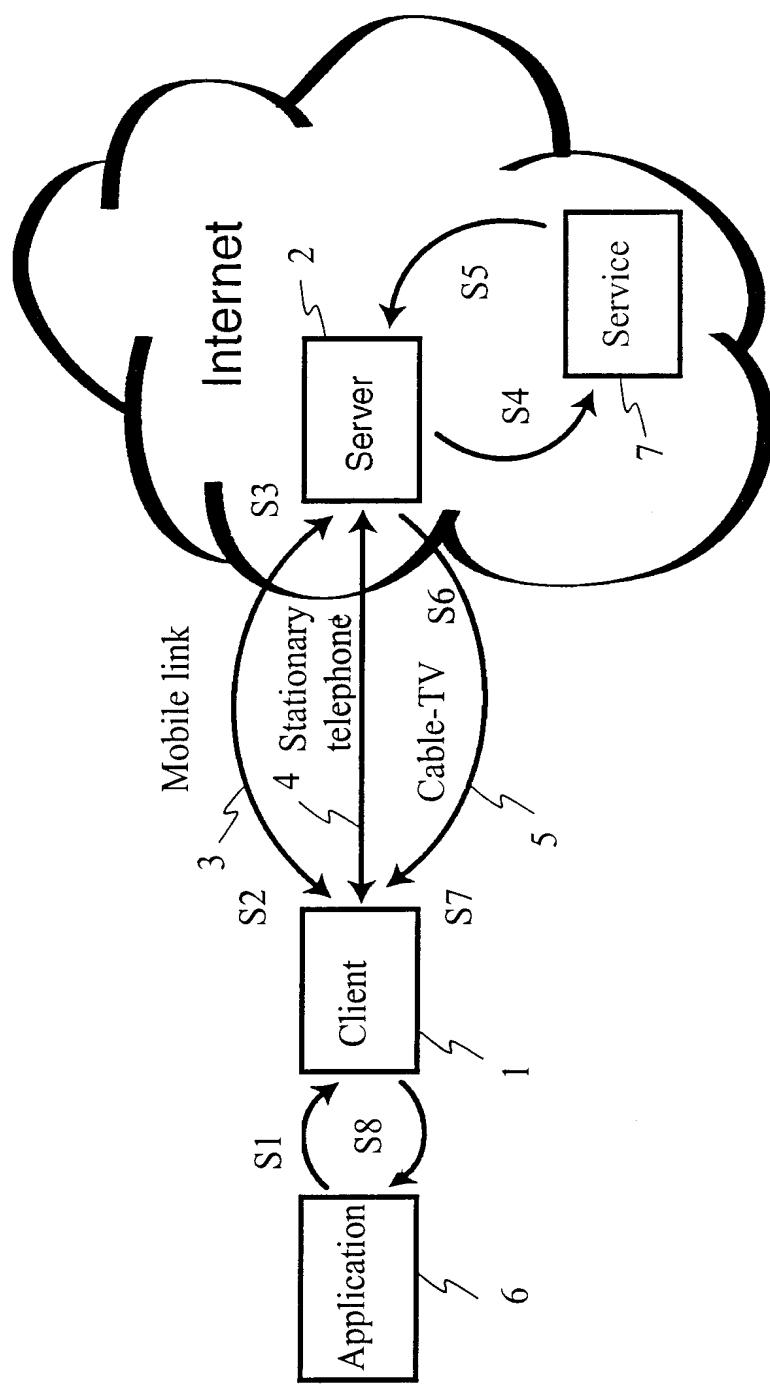


Fig. 1

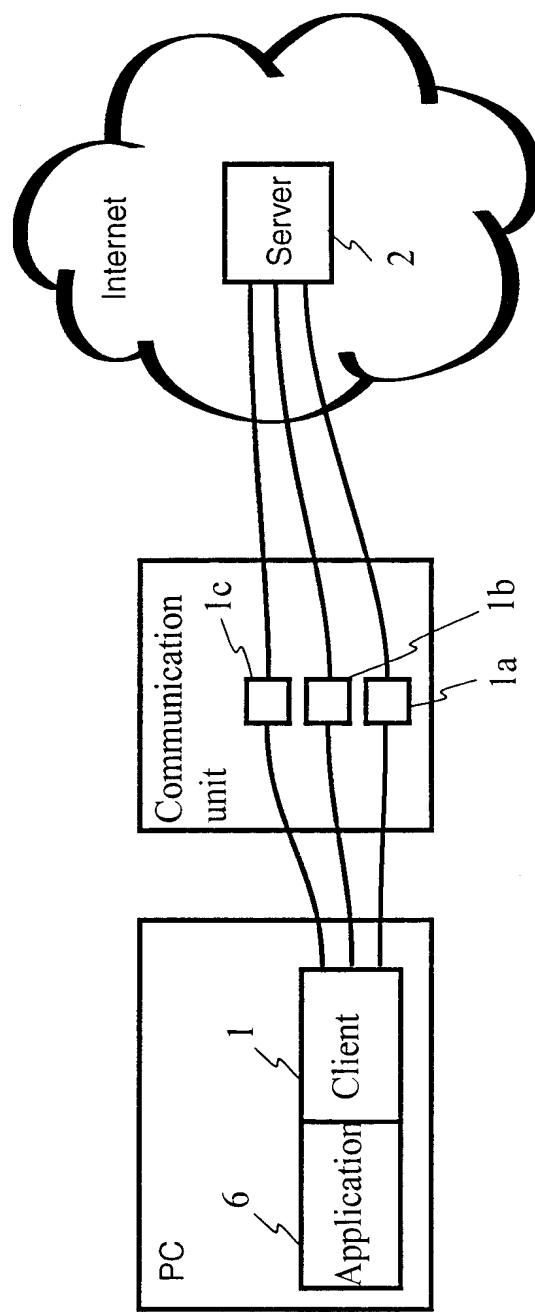


Fig. 2

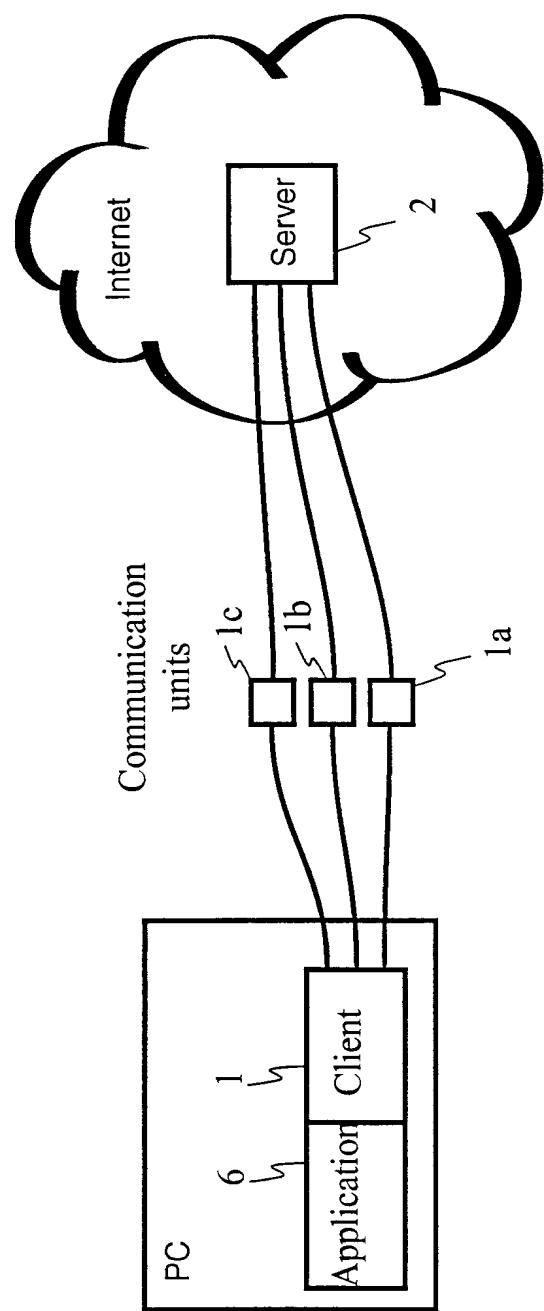


Fig. 3

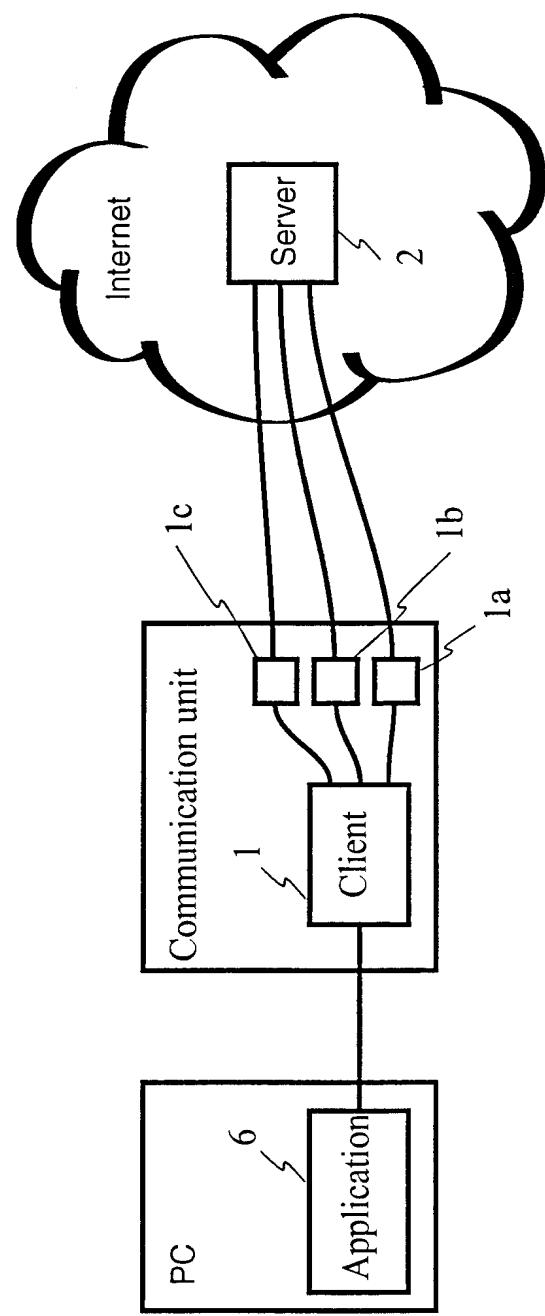


Fig. 4

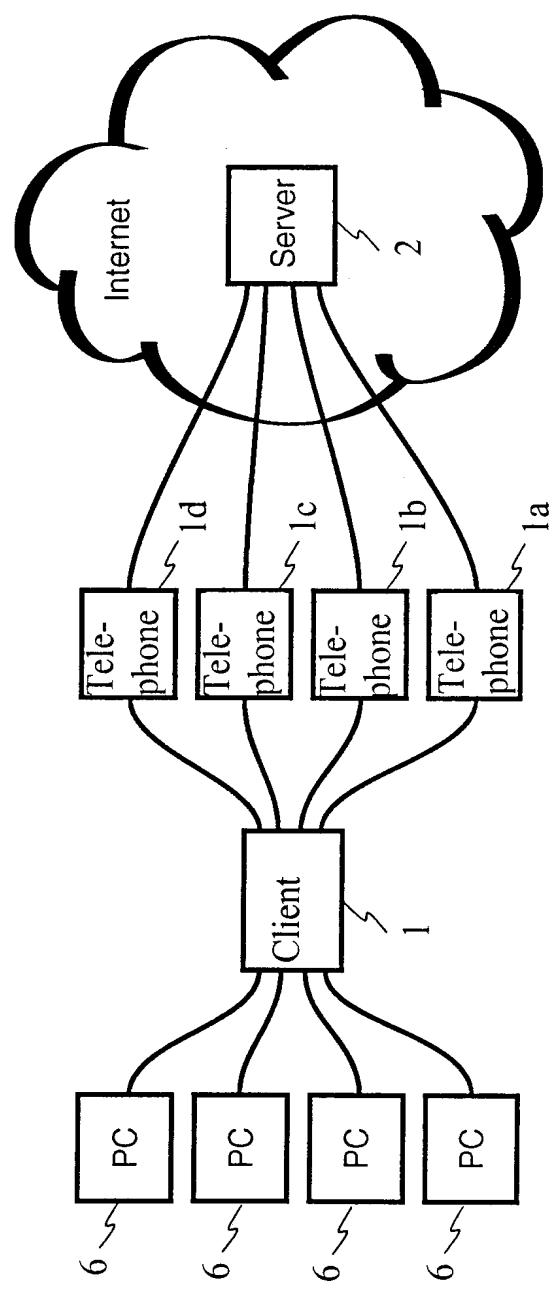


Fig. 5

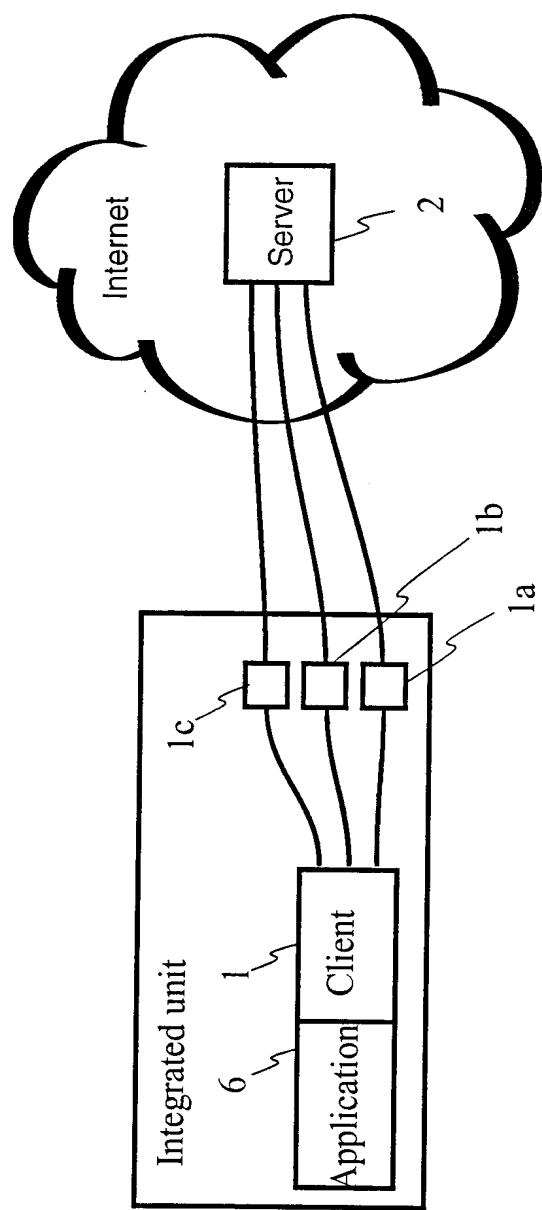


Fig. 6

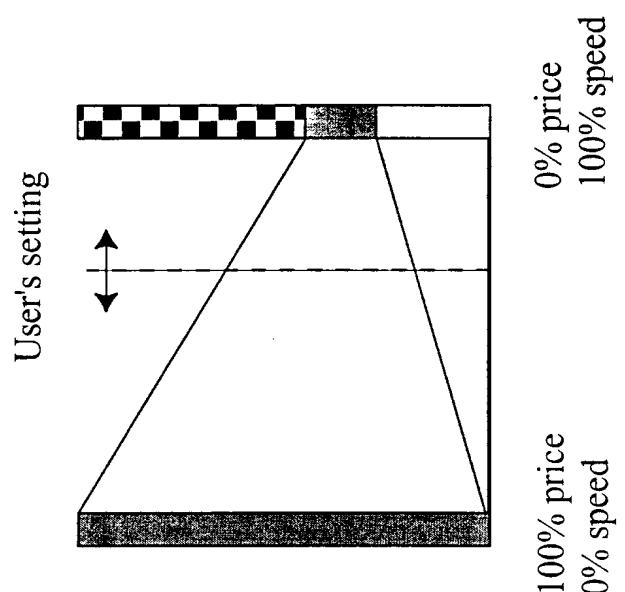


Fig. 7

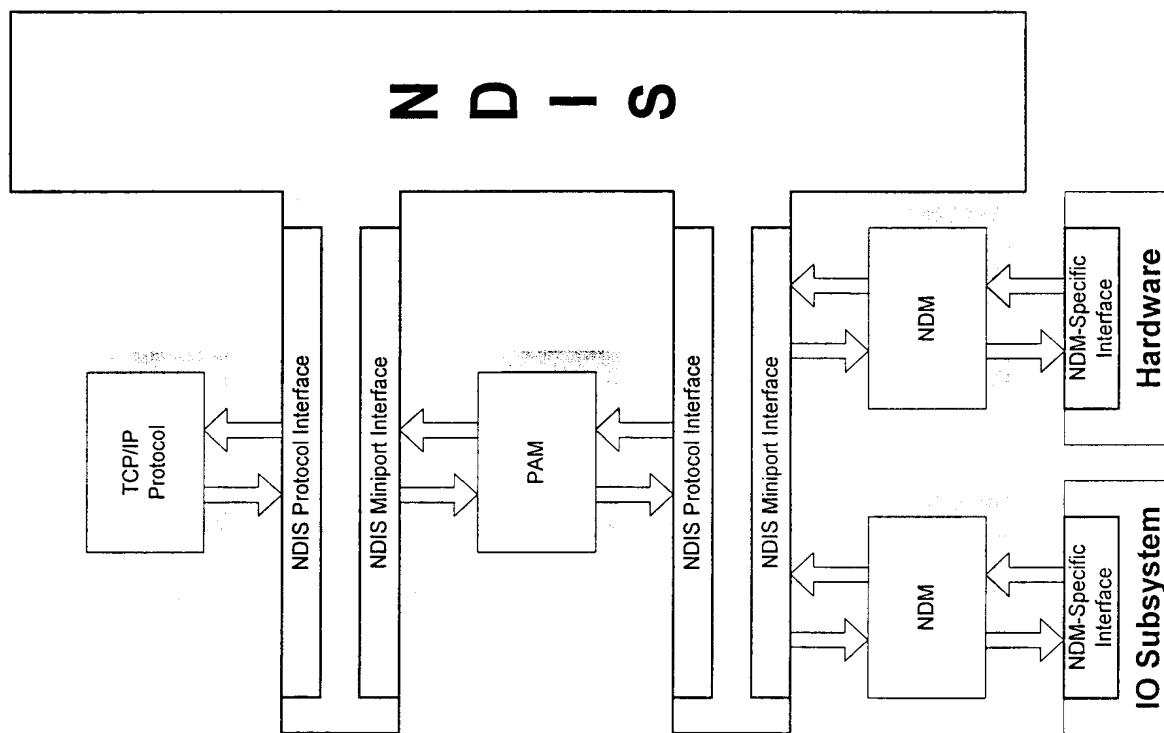


Fig. 8

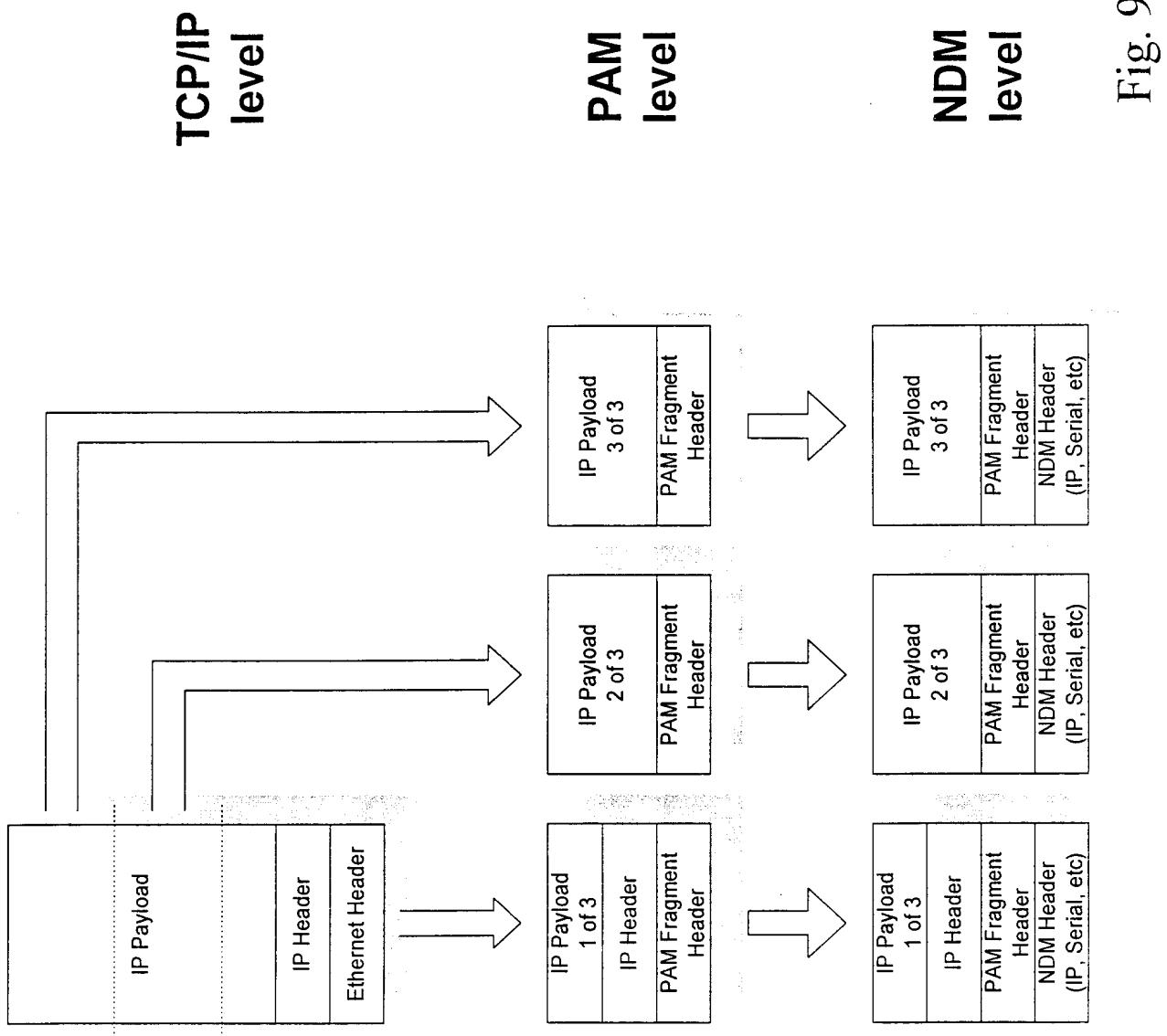
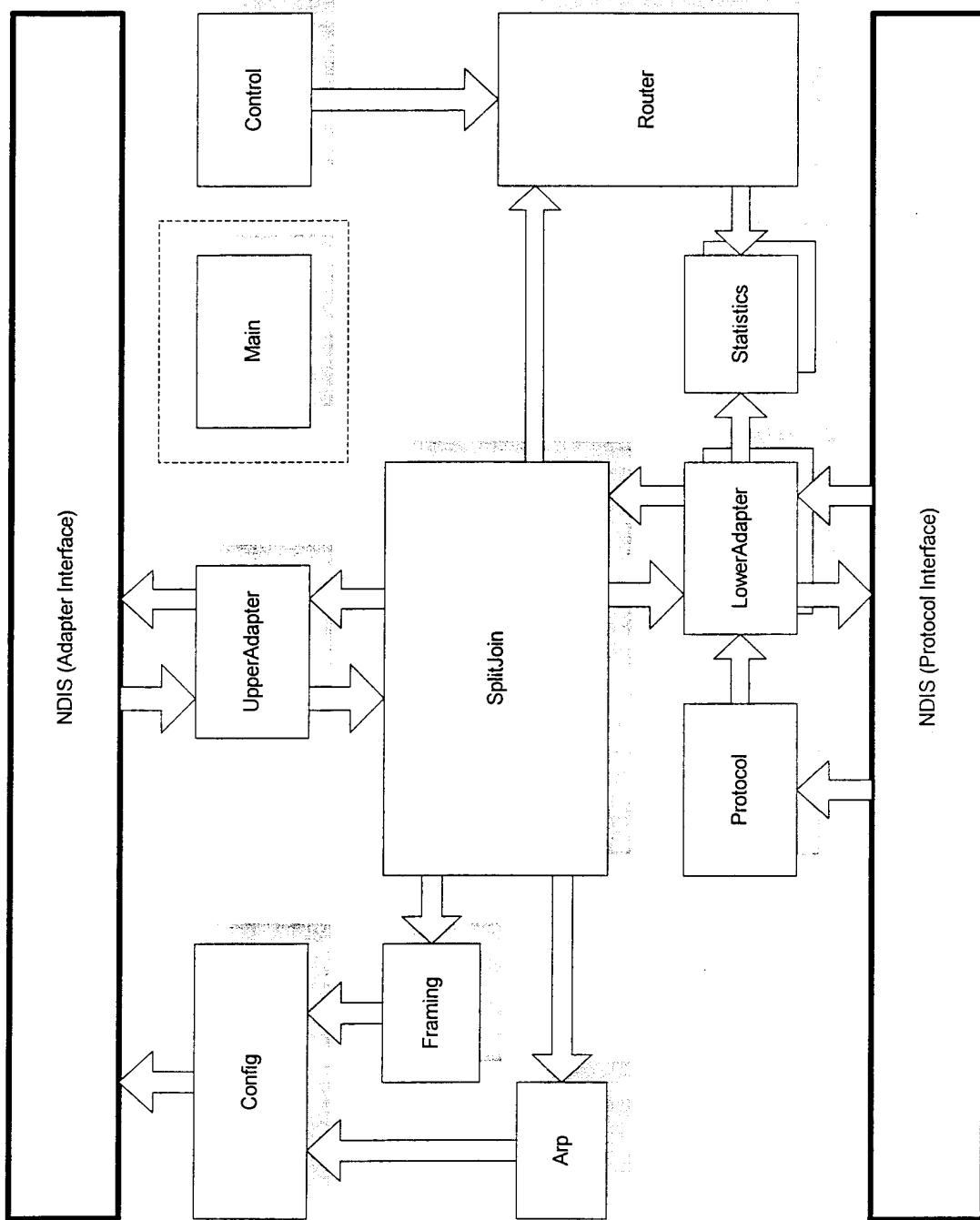


Fig. 9

Fig. 10



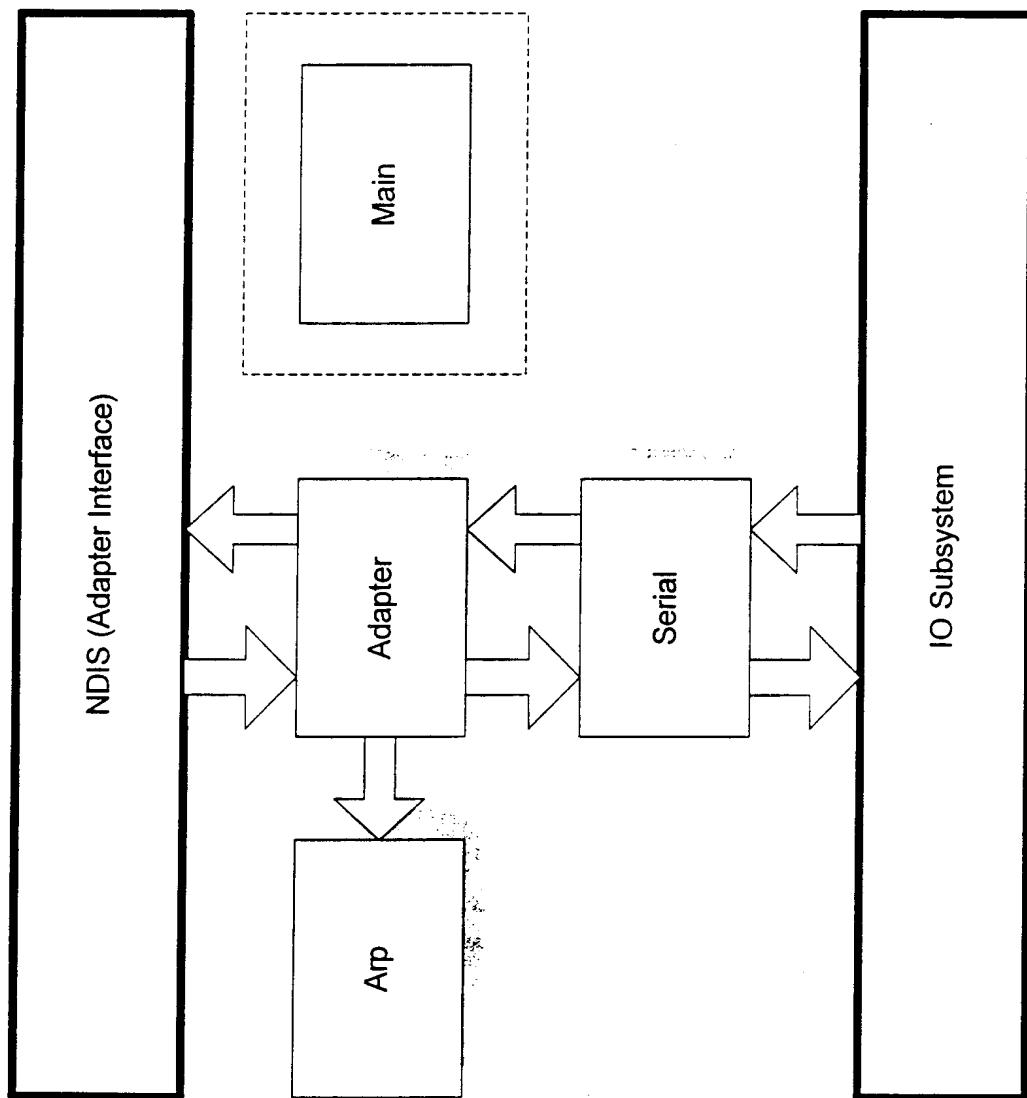


Fig. 11

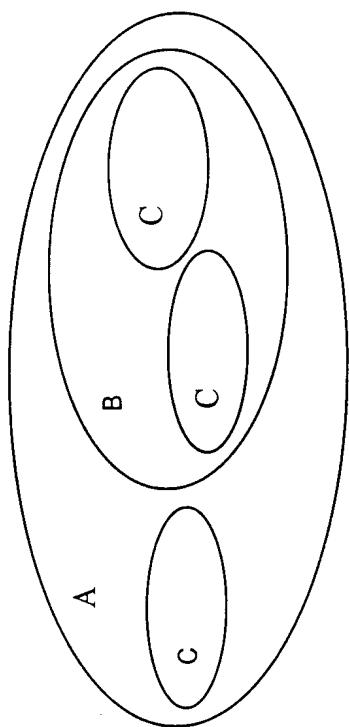
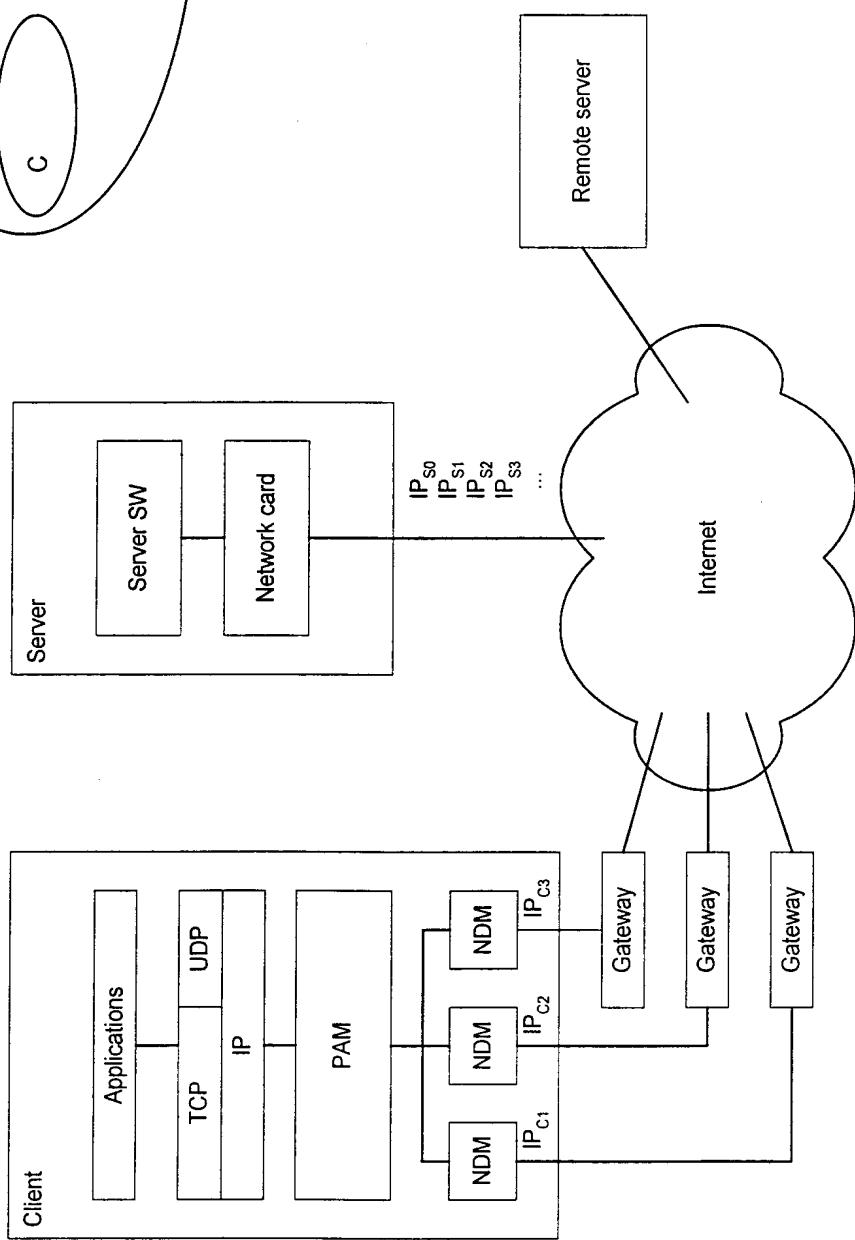


Fig. 12



INTERNATIONAL SEARCH REPORT

International application No.

PCT/SE 00/00883

A. CLASSIFICATION OF SUBJECT MATTER

IPC7: H04L 12/56, H04Q 11/04, H04L 29/06

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC7: H04L, H04Q, H04F, H04M

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SE,DK,FI,NO classes as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	EP 0855820 A2 (INTERNATIONAL BUSINESS MACHINES CORPORATION), 29 July 1998 (29.07.98), page 3, line 21 - line 53; page 4, line 13 - page 5, line 14, claims 1-9, abstract --	1-17
A	WO 9819439 A1 (CONNECT ONE, INC.), 7 May 1998 (07.05.98), page 5, line 25 - page 7, line 15; page 7, line 25 - page 9, line 5, figure 1, claims 1-24 --	1-17
A	EP 0615391 A1 (ALCATEL N.V.), 14 Sept 1994 (14.09.94), column 16, line 17 - line 29, claims 1-16 --	1-17

Further documents are listed in the continuation of Box C.

See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"I" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

Date of mailing of the international search report

19 Sept 2000

29 -09- 2000

Name and mailing address of the ISA/
Swedish Patent Office
Box 5055, S-102 42 STOCKHOLM
Facsimile No. + 46 8 666 02 86

Authorized officer

Roger Bou Faisal/LR
Telephone No. + 46 8 782 25 00

INTERNATIONAL SEARCH REPORT

International application No.

PCT/SE 00/00883

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
E	WO 9943133 A2 (NOKIA TELECOMMUNICATIONS OY), 26 August 1999 (26.08.99), see the whole document --	1-17
A	RICHARD W. WATSON ET AL: "The Parallel I/O Architecture of the High-Performance Storage System (HPSS)", Fourteenth IEEE Symposium on Mass Storage Systems; IEEE, Volume, 1995 --	1-17
A	YARON RESHEF ET AL:"Reliable Multilink Protocols", IEEE Transactions on Communications, Vol 41, No 12, December 1993 -- -----	1-17

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/SE 00/00883

Patent document cited in search report		Publication date		Patent family member(s)		Publication date
EP	0855820	A2	29/07/98	JP US	10190738 A 5956341 A	21/07/98 21/09/99
WO	9819439	A1	07/05/98	BR CN EP US	9712399 A 1235729 A 0931408 A 6016307 A	31/08/99 17/11/99 28/07/99 18/01/00
EP	0615391	A1	14/09/94	AU AU CA CN FI JP US	667763 B 5514594 A 2118559 A 1099545 A 941093 A 6350520 A 5504935 A	04/04/96 15/09/94 10/09/94 01/03/95 10/09/94 22/12/94 02/04/96
WO	9943133	A2	26/08/99	AU FI FI	2282399 A 3652 U 980294 A,V	06/09/99 09/10/98 10/08/99